# The First Phylogenetic Analysis of Palpigradi (Arachnida)—The Most Enigmatic Arthropod Order

# The first phylogenetic analysis of Palpigradi (Arachnida)—the most enigmatic arthropod order

*Gonzalo Giribet[A,K], Erin McIntyre[A], Erhard Christian[B], Luis Espinasa[C], Rodrigo L. Ferreira[D], Óscar F. Francke[E], Mark S. Harvey[F], Marco Isaia[G], Ľubomīr Kováč[H], Lynn McCutchen[I], Maysa F. V. R. Souza[D] and Maja Zagmajster[J]*

[A]Museum of Comparative Zoology, Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA.

[B]Institut für Zoologie, Universität für Bodenkultur, Gregor-Mendel-Straße 33, 1180 Wien, Austria.

[C]School of Science, Marist College, 3399 North Road, Poughkeepsie, New York, USA.

[D] Centro de Estudos em Biologia Subterrânea, Departamento de Biologia, Universidade Federal de Lavras, Lavras, MG. CEP 37200-000, Brazil.

[E]Colección Nacional de Arácnidos, Instituto de Biologia, UNAM, Apartado Postal 70-153, C. P. 04510, Mexico, D. F., Mexico.

[F]Department of Terrestrial Zoology, Western Australian Museum, Locked Bag 49, Welshpool DC, WA 6986, Australia.

[G]Dipartimento di Scienze della Vita e Biologia dei Sistemi, Università di Torino, Via Accademia Albertina 13, 10123 Torino, Italy.

[H]Department of Zoology, Institute of Biology and Ecology, Faculty of Science, P. J. Šafárik University, Moyzesova 11, 040 01 Košice, Slovakia.

[I]Department of Biology, Kilgore College, 1100 Broadway, Kilgore, TX 75662, USA.

[J]SubBioLab, Department of Biology, Biotechnical Faculty, University of Ljubljana, Večna pot 111, SI-1000 Ljubljana, Slovenia.

[K]Corresponding author. Email: ggiribet@g.harvard.edu

**Abstract.** Palpigradi are a poorly understood group of delicate arachnids, often found in caves or other subterranean habitats. Concomitantly, they have been neglected from a phylogenetic point of view. Here we present the first molecular phylogeny of palpigrades based on specimens collected in different subterranean habitats, both endogean (soil) and hypogean (caves), from Australia, Africa, Europe, South America and North America. Analyses of two nuclear ribosomal genes and COI under an array of methods and homology schemes found monophyly of Palpigradi, Eukoeneniidae, and a division of Eukoeneniidae into four main clades, three of which include samples from multiple continents. This supports either ancient vicariance or long-range dispersal, two alternatives we cannot distinguish with the data at hand. In addition, we show that our results are robust to homology scheme and analytical method, encouraging further use of the markers employed in this study to continue drawing a broader picture of palpigrade relationships.

**Additional keywords:** Arachnida, micro-whip scorpions, palpigrades, speleobiology, biogeography.

## Introduction

The arachnid order Palpigradi (micro-whip scorpions or palpigrades) is one of the smallest, rarest and most neglected groups of terrestrial arthropods, and one of the last arachnid orders to be discovered—it was first reported only in 1885 (Grassi and Calandruccio 1885). The first photographs of living palpigrades did not appear published until the first decade of the 21st century (Kováč *et al.* 2002; Beccaloni 2009). Additionally, only a handful of DNA sequence data are available in GenBank; with only 64 sequences, 56 are for *Prokoenenia wheeleri* (Rucker, 1901), a species that was part of a multi-gene phylogeny of arthropods (Regier *et al.* 2010), while the remaining eight sequences are unidentified specimens from three studies on chelicerate phylogenetics (Giribet *et al.* 2002; Pepato *et al.* 2010; Arabi *et al.* 2012). Contrary to this, one can find more DNA sequences for other small arachnid orders in GenBank: 105 for Uropygi, 200 for Schizomida, 200 for Ricinulei, 251 for Amblypygi, and 502 for Pseudoscorpiones, [checked on October 25th, 2013]. In addition, there are only 2 sequences available in the Barcode of Life website (http://www.barcodinglife.org).

Palpigrades are delicate animals that walk sensing the substrate with what seems a nervous behaviour of the first pair of walking legs, and use their unmodified palps for walking, unlike all other arachnids (Fig. 1). While moving, most palpigrades keep the flagellum upward, moving it laterally. Accordingly, it is possible that the uplifted flagellum is associated with perception of the environment (Ferreira and Souza 2012). These small, depigmented and highly translucent arachnids range in size from 0.65 mm in *Eukoenenia grassii* (Hansen, 1901) to 2.4 mm in the "giant" *E. draco* (Peyerimhoff, 1906) from caves on the island of Majorca (Mayoral and Barranco 2013). *Eukoenenia spelaea* (Peyerimhoff, 1902) from Slovakia has recently been reported to feed on heterotrophic Cyanobacteria (Smrž *et al.* 2013). The mode of sperm transfer in these arachnids remains unknown.

The living members of the order are currently divided in two families, Eukoeneniidae Petrunkevitch, 1955, with 4 genera and 85 named species, and Prokoeneniidae Condé, 1996, with 2 genera and 7 named species (Harvey 2002; Prendini 2011; Souza and Ferreira 2013). Eukoeneniidae includes the genera *Allokoenenia* Silvestri, 1913 (1 sp. from West Africa), *Eukoenenia* Börner, 1901 (71 spp., on all continents under tropical and subtropical climate; in temperate regions predominantly in caves), *Koeneniodes* Silvestri, 1913 (8 Palaeotropical spp.) and *Leptokoenenia* Condé, 1965 (5 spp. in the Afrotropical, Neotropical and Palearctic regions).

Prokoeneniidae includes the genera *Prokoenenia* Börner, 1901 (6 spp. in the Nearctic, Neotropical and Oriental regions) and *Triadokoenenia* Condé, 1991 (1 sp. from Madagascar). Further unnamed new species are known to us from various parts of the world.

The position of Palpigradi among the arachnid orders remains highly debated. The largest set of data analysed to date places them as the sister group to Acariformes mites in a basal position within arachnids, although without support (Regier *et al.* 2010). The most recent morphological cladistic analysis of arachnid relationships leaves them mostly unresolved among the clades Stomothecata, Haplocnemata, Pantetrapulmonata, and Acaromorpha (Shultz 2007). Earlier studies combining morphology and a small set of molecular data placed Palpigradi as the sister group of Ricinulei + Tetrapulmonata or as sister to Pycnogonida when fossils were considered, although again, without significant clade support (Giribet *et al.* 2002); as sister to a clade including Acari and Solifugae, based on the same two markers used in earlier studies (Pepato *et al.* 2010); or in an unresolved position within arachnids (Arabi *et al.* 2012). Even less is known about the internal relationships of the group, since no published study—molecular or morphological—has yet incorporated information for more than one palpigrade species, and only one unpublished masters thesis has explored palpigrade relationships cladistically, using morphology (Montaño Moreno 2008).

To bridge this important gap in the knowledge of this arachnid order, although acknowledging the difficulties in sampling and identification of these elusive animals, we obtained samples for as many species of palpigrades as possible and from as many localities as possible with the aim to obtain molecular DNA sequence data to generate a first hypothesis of internal palpigrade relationships.

## Materials and Methods

### *Taxon sampling*

Palpigrades are difficult to obtain and identify, and success of field sampling differed among regions included in the study. In Western Australia, many samples were collected indirectly in caves and bore holes. In Brazil and Europe, they can be abundant in caves, where fresh specimens have recently become available for inclusion in molecular studies. Additional samples

were from soil samples in Australia, Italy and the USA. In addition to fresh material collected for this study, older specimens were used, especially from the diverse cave systems in Brazil, where several new species have been recently described (Souza and Ferreira 2010; Ferreira *et al.* 2011; Souza and Ferreira 2011a; Souza and Ferreira 2011b; Souza and Ferreira 2012a; Souza and Ferreira 2012b). While a recently collected specimen of *Eukoenenia ferratilis* Souza & Ferreira, 2011 amplified well for some of the studied markers, none of the six specimens of *Allokoenenia* spp. and the two specimens of *Leptokoenenia* sp. collected from the caves yielded workable DNA. We also obtained a relatively large collection of specimens from the Western Australian bore holes from Barrow Island and the Pilbara, but these were collected from litter traps and many specimens did not amplify or only yielded some amplicons. Some of these specimens are probably related to the Western Australian endemic *E. guzikae* Barranco & Harvey, 2008, but unrelated to the more widespread species *E. mirabilis* (Grassi & Calandruccio, 1885), also found in Western Australia (Harvey *et al.* 2006; Barranco and Harvey 2008). A single specimen of *Prokoenenia wheeleri* was obtained from the Austin area (Texas, USA), but amplified well for all fragments attempted. In addition, we obtained samples of *Eukoenenia mirabilis* from Italy (Christian *et al.* 2010) and Australia (Harvey *et al.* 2006), *E. spelaea* (Peyerimhoff, 1902) from multiple localities in Slovenia and Slovakia (Kováč *et al.* 2002; Zagmajster and Kováč 2006; Král *et al.* 2008). Italian samples also include *E. bonadonai* Condé, 1979 and *E. strinatii* Condé, 1977, collected in caves. We also included specimens from multiple localities from the hanseni-chilanga group of *Eukoenenia* from Mexico and the USA (Montaño-Moreno 2012). Additional specimens come from Mexican caves and South Africa. Details on collecting localities are available in Table 1 and in MCZBASE (http://mczbase.mcz.harvard.edu/SpecimenSearch.cfm). Vouchers or additional specimens are deposited in the Museum of Comparative Zoology, Harvard University (MCZ), and in the Western Australian Museum (WAM).

We included three species available in GenBank, one from South Africa sequenced by Giribet *et al.* (2002), one from Brazil from Pepato *et al.* (2010), and one of unknown origin published by Arabi *et al.* (2012). Here we added sequences from an additional South African specimen from the same collection of that from Giribet *et al.* (2002), and a specimen of *E. ferratilis* from Brazil, which was identical to the specimen reported by Pepato *et al.* (2010) as *Eukoenenia* sp., and to which we refer to as *E.* cf. *ferratilis* in the present study. Outgroup taxa were selected from GenBank (Table 2), mostly from previous studies on arthropod or arachnid phylogeny using nuclear ribosomal genes (Giribet *et al.* 2002; Mallatt and Giribet 2006).

133

134 *Molecular methods*

135 Although we attempted to amplify and sequence five molecular markers typically used in other
136 analyses of arachnid systematics (e.g., Dimitrov *et al.* 2012; Giribet *et al.* 2012), the
137 mitochondrial 16S rRNA gene only amplified for *Prokoenenia wheeleri* and the nuclear protein-
138 encoding gene histone H3, although amplified for several samples, did not produce clean reads.
139 We thus restricted our study to the two broadly available nuclear ribosomal genes, the
140 complete 18S rRNA and ca. 2.2 Kb of 28S rRNA, and the mitochondrial protein-encoding
141 cytochrome *c* oxidase subunit I (COI hereafter) (as in Murienne *et al.* 2008), although the latter
142 gene only amplified for about a third of the specimens (Table 1). For two of the bore-hole
143 Western Australian specimens, poorly preserved, only the middle amplicon of 28S rRNA worked.

144 Total DNA was extracted from whole specimens or from the opisthosomal region using
145 Qiagen's DNEasy® tissue kit (Valencia, CA, USA). Although we were aiming to preserve the
146 digested carcass as a morphological voucher, it was completely digested and not recoverable.
147 Purified genomic DNA was used as a template for Polymerase chain reactions (PCR)
148 amplification. PCR, visualization by agarose gel electrophoresis, and direct sequencing were
149 conducted for most specimens as described in earlier work, e.g., Edgecombe and Giribet (2009).
150 Chromatograms obtained from the automatic sequencer were read and sequences assembled
151 using the sequence editing software Sequencher™ (Gene Codes Corporation, Ann Arbor, MI,
152 USA). Sequence data were edited in MacGDE (Linton 2005). The three genes were analysed as
153 follows:

154 18S rRNA: This marker was amplified in three amplicons (*a, b, c*), as in previous studies
155 (Edgecombe and Giribet 2009; Giribet *et al.* 2010; Giribet *et al.* 2012). In the present study we
156 include 27 palpigrade specimens plus 8 outgroups, for a total of 1760-1771 bp per complete
157 sequence (up to 1805 bp for one of the outgroups). From the 27 palpigrade sequences all but
158 three were complete; *E. spelaea* is missing fragment *a* and the sample of *Eukoenenia* from South
159 Africa (DNA100456.2) is missing fragment *b*. For the direct optimization analyses the three
160 amplicons were treated as a single input file, containing 23 sequences, and divided into six
161 fragments. The three amplicons were concatenated for the static alignment analyses.

28S rRNA: This nuclear gene was amplified in three amplicons (*a, b, c*), as described in Giribet and Shear (2010). The data set includes 29 palpigrade specimens plus 8 outgroups, for a total of 2,150 to 2,204 bp, with some length variation among species. These three fragments correspond to primer pairs 28S rd1a—28D rd4b, 28Sa—28S rd5b, and 28S rd4.8a—28S rd7b1. Some of the published sequences were amplified with a shorter fragment *b*, generated with primers 28Sa—28Sb (Whiting *et al.* 1997), and therefore fragment *b* was divided into fragments *b1* and *b2* to accommodate these two amplicons. Fragment *a* was available for 22 palpigrades and divided into three fragments, fragment *b* for 29 palpigrades and three fragments, and fragment *c* for 25 palpigrades and analysed as a single fragment. These were treated as three different amplicons for the dynamic homology analyses, but aligned together for the static homology approaches.

COI: This widely used mitochondrial marker amplified for ten palpigrade terminals in a single amplicon using primers LCO—HCO, showing no length variation (654 bp analysed), plus one was available in GenBank. COI did not amplify for many individuals, perhaps due to major changes in this marker, as evidenced by the deletion of one amino acid with respect to the outgroups. Five outgroup sequences were obtained from GenBank, but these were 3 bp longer in all cases except for the pseudoscorpion. It was analysed as a single fragment; not pre-aligned due to the length difference with some outgroups.

*Phylogenetic analyses*

Parsimony analyses were based on a direct optimization (DO) approach (Wheeler 1996) using POY v. 5.0 (Varón *et al.* 2012). Tree searches were performed using the timed search function in POY, i.e., multiple cycles of (a) building Wagner trees, (b) subtree pruning and regrafting (SPR), and (c) tree bisection and reconnection (TBR), (d) ratcheting (Nixon 1999), and (e) tree-fusing (Goloboff 1999, 2002) [command: `search (max_time:00:01:00, min_time:00:00:10, hits:20, memory:gb:2)`]. For the individual partitions, timed searches of 1 hour were run on 4 processors under six parameter sets, as in Giribet *et al.* (2012) (see Table 3). For the combined analysis of the three markers we started with the same search strategy, giving the 28S rRNA trees as input—as these contained all the taxa in the combined data set—, and the resulting trees were given as input for a second round of analyses (sensitivity analysis tree fusing; SATF),

as described by Giribet (2007), and continued until the tree lengths stabilised (Giribet *et al.* 2012). The optimal parameter set was estimated using the modified $_{W}$ILD metrics (Wheeler 1995; Sharma *et al.* 2011), as a proxy for the parameter set that minimizes overall incongruence among data partitions (Table 4). Nodal support for the optimal parameter set was estimated via jackknifing (250 replicates) with a probability of deletion of $e^{-1}$ (Farris *et al.* 1996) using `auto_sequence_partition`, as discussed in earlier work (Giribet *et al.* 2012).

Maximum likelihood (ML) analyses were conducted on static multiple sequence alignments (MSA) inferred in MUSCLE v. 3.6 (Edgar 2004) through the EMBL-EBI server (http://www.ebi.ac.uk/Tools/msa/muscle/). We also used an implied alignment (IA) generated in POY (Wheeler 2003; Giribet 2005) for subsequent analyses based on static alignments, as recently explored by Giribet and Edgecombe (2013b) for a centipede data set. The MUSCLE alignments were conducted for each gene independently. The IA and MSA therefore were based on the same data (see length for each gene in Table 5). In order to evaluate the impact of the hypervariable regions in the data set, MSAs and IAs were subsequently trimmed with Gblocks v. 0.91b (Castresana 2000; Talavera and Castresana 2007) to cull positions of ambiguous homology (see length for each trimmed gene in Table 5). In the case of 28S, fragments *a* and *bc* were Gblocked separately, due to the larger proportion of missing data in the *a* fragment, which otherwise would be deleted from the final 28S alignment. These data sets are thus based on different data from their original sources and from each other, but the remaining data use the same homology scheme as the source. Data sets were concatenated with SequenceMatrix (Vaidya *et al.* 2011).

Maximum likelihood analyses were conducted using RAxML ver. 7.2.7 (Stamatakis *et al.* 2008b) in the CIPRES server (Miller *et al.* 2010). For the searches, a unique General Time Reversible (GTR) model of sequence evolution with corrections for a discrete gamma distribution (GTR + $\Gamma$) was specified for each data partition, and 100 independent searches were conducted. Nodal support was estimated via the rapid bootstrap algorithm (1000 replicates) using the GTR-CAT model (Stamatakis *et al.* 2008a). Bootstrap resampling frequencies were thereafter mapped onto the optimal tree from the independent searches.

In total we analysed five data sets accounting for different optimality criteria, homology schemes, and/or amount of data, as follows:

- Analysis 1. Direct optimization/dynamic homology under parsimony (full sensitivity analysis of 6 parameter sets) analysed in POY
- Analysis 2. Static homology from the implied alignment for the optimal parameter set under ML (analysed in RAxML)
- Analysis 3. Static homology from the implied alignment for the optimal parameter set trimmed with Gblocks under ML (analysed in RAxML)
- Analysis 4. Static homology based on MUSCLE multiple sequence alignment (analysed in RAxML)
- Analysis 5. Static homology based on MUSCLE/Gblocks (analysed in RAxML)

**Results and Discussion**

All phylogenetic analyses yielded very similar results with respect to the ingroup relationships, while the outgroup relationships were incongruent from analysis to analysis and unsupported for the most part (Figs. 2 and 3). The latter was expected given the small amount of data and outgroup taxa and the poor resolution in deep arachnid relationships in other studies (e.g., Wheeler and Hayashi 1998; Giribet *et al.* 2002; Pepato *et al.* 2010; Regier *et al.* 2010). The optimal parameter set under parsimony direct optimization was *3211* (where indel opening costs 3, indel extension 1, transversions cost 2 and transitions cost 1; $_{W}$ILD = 0.00913), with a cost of 10,408 weighted steps (Fig. 2). Nearly all examined parameter sets concurred on the topology of the optimal parameter set, with the exception of *Eukoenenia spelaea* IZ-19346 from Slovenia, and the resolution of one of the *Eukoenenia* clades (see below). Likewise, the analyses of the four data sets analysed under maximum likelihood were nearly identical, except for some of the shallowest relationships. One of these trees, the one for the multiple sequence alignment trimmed with Gblocks—the one that could be potentially the most different from the POY analysis—is presented in Fig. 3, and it is virtually identical to the direct optimization tree. From the 10 nodes depicted in Fig. 2 summarizing the six direct optimization and the four maximum likelihood analyses, 5 were recovered in all analyses. Support values for these five nodes is high for most analyses (jackknife values are lower by definition), with the exception of clades III and IV in the DO analysis. Basically, nearly all analyses concur on the overall topology of the palpigrade tree.

All analyses show a basal dichotomy between *Prokoenenia wheeleri* (the only Prokoeneniidae represented in our analyses) and the remaining samples, which we consider as *Eukoenenia* for further discussion—even if some samples from GenBank or from the Australian boreholes were not identified. *Eukoenenia* is divided into four main clades, indicated in Figures 2 and 3. Clade I includes *E. florenciae* from Slovakia, Brazil, and unidentified specimens probably belonging to the same species from the USA and Mexico, and another species from a cave in Guerrero, Mexico (IZ-128499). Clade II includes *E. spelaea* and *E. s. hauseri* Condé, 1974 from Slovenia and Slovakia, and several additional samples from Slovenia and Italy, including *E. strinatii*, *E. bonadonai* and *E. austriaca* (Hansen, 1926); *E. spelaea* IZ-19346 from Slovenia clusters with these species in some analyses, but not all (Fig. 2). Clade III includes *E. ferratilis* from Brazil, the specimens from the Australian bore holes, and an undescribed species from Brazil (IZ-19345). Clade IV includes *E. mirabilis* from Australia and Italy, and unidentified specimens from South Africa, plus a specimen from a cave in Chiapas, Mexico (IZ-136274) and a GenBank specimen (JA-2011) of unknown origin. Clades I and II are supported in all analyses; Clade III is supported in all analyses except for the DO analysis under parameter set 211; Clade IV is unsupported in the ML analysis of the trimmed MSA. *Eukoenenia spelaea* IZ-19346 appears as the sister group to Clade II under 4 analytical parameter sets in DO and in the untrimmed ML analyses, both for the IA and for the MSA. The *E. florenciae* clade (Clade I) always forms the sister group of the *E. spelaea* clade (Clade II), although *E. spelaea* IZ-19346 sometimes forms the sister group of the *E. florenciae* clade. While the *E. ferratilis* clade (Clade III) often forms the sister group to the *E. mirabilis* clade (Clade IV) (Figs. 2, 3), and is well supported in the probabilistic analyses (97 to 100% bootstrap support, depending on the analysis), under some parameter sets Clade III is sister to the *E. spelaea—E. florenciae* clade (parameter sets 111, 211, 221, 3221).

Irrespective of these small differences, our analyses show high congruence between alternative methods (parsimony and maximum likelihood) based on identical raw data with different homology schemes (implied alignments versus multiple sequence alignments), or different data sets (trimmed implied alignments and trimmed multiple sequence alignments). There are very few cases with such consistency across weighting schemes, homology schemes, and methodologies, but a recent case was documented for scutigeromorph centipedes (Giribet and Edgecombe 2013b). In that case, the fossil record and denser sampling allowed for accurate molecular dating and analyses of diversification of lineages through time, and it was suggested

284 that the congruence across analyses was due to constant rates of diversification through more
285 than 400 million years of evolution in the group. We can only guess this for palpigrades, as the
286 fossil record for this group is rare, and a single Pliocene specimen is known (Rowland and Sissom
287 1980; Delclòs *et al.* 2008; Dunlop 2010), although the group must be much older in origin (see
288 for example Giribet and Edgecombe 2013a).

289 Phylogenetic analysis of the three molecular markers combined and for all analyses
290 performed resolves into Prokoeneniidae (although represented by a single species) and
291 Eukoeneniidae, supporting the monophyly of Eukoeneniidae—palpigrades without sternal
292 opisthosomal vesicles (Condé 1996). We were, however, unable to obtain samples of
293 *Triadokoenenia* or of additional *Prokoenenia* species, thus not being able to test the taxon
294 Prokoeneniidae. Within Eukoeneniidae, the four main clades discussed above are supported in
295 nearly all analyses. But species identifications in palpigrades do not seem straightforward.
296 Within Clade I, the specimens of *Eukoenenia* from Texas (USA), the Mexican state of Yucatán, *E.*
297 cf. *florenciae* from Brazil and *E. florenciae* from Slovakia show nearly identical COI sequences
298 and identical nuclear ribosomal RNA sequences, suggesting that they may be conspecific (see
299 Edgecombe and Giribet 2008; Vélez *et al.* 2012). In contrast, Clade II includes three lineages of
300 the morphospecies *E. spelaea*. From these, two samples identified as *E. spelaea* and *E. spelaea*
301 *hauseri* from Slovenia appear identical for the nuclear ribosomal genes (but did not amplify for
302 COI).

303 Clade III includes the Western Australian samples and *Eukoenenia ferratilis* from the Iron
304 caves of Minas Gerais (Brazil). Difficulties in amplifying the Australian samples and the lack of
305 COI information for any of the members of the clade precludes us from understanding genetic
306 variability within this clade of geographically distant species (both between the continents, but
307 also among the Western Australian localities), although most analyses consistently resolve this
308 clade of six individuals with reciprocal monophyly of the two geographic regions.

309 Clade IV, although with less support than the other three clades, includes the sample of
310 unknown provenance sequenced by Arabi *et al.* (2012), a specimen from caves in Chiapas, and
311 the cosmopolitan *E. mirabilis*, including two specimens from Italy (identical for all markers) and
312 two putative members of this species from South Africa plus a sample of *E. mirabilis* from
313 Australia. While *E. mirabilis* has been suggested to be a synanthropic species originating in the
314 Mediterranean region with recent introductions to South Africa, Australia, Chile and

315 Madagascar (Harvey *et al.* 2006), our limited data suggest a close relationship between one of
316 the South African samples and the Australian specimen, even in the absence of COI data, and
317 therefore suggesting changes in the nuclear ribosomal genes with respect to the Italian sample.
318 Further study of Gondwanan *E. mirabilis* and addition of circum-Mediterranean samples should
319 be undertaken to bring this matter to conclusion.

320 Given the sampling of this study it is still early to make any firm conclusions about
321 palpigrade relationships. We were not able to test for the monophyly of Prokoeneniidae, and
322 monophyly of *Eukoenenia* is not thoroughly tested either. Attempts to sequence *Allokoenenia*
323 and *Leptokoenenia* were unsuccessful, and we were unable to obtain specimens of the
324 Palaeotropical *Koeneniodes* and *Triadokoenenia*. Few studies have looked at variation among
325 palpigrade species, but Král *et al.* (2008) investigated the karyotypes of *E. spelaea* from Slovakia
326 and *E. mirabilis*, which appear in different clades in our study (Clades II and IV, respectively).
327 However, the karyotypes of both species showed no variation, both consisting of a low number
328 of tiny chromosomes that decrease gradually in size and a lack of morphologically differentiated
329 sex chromosomes, suggesting that molecular data may be more informative than karyotypic
330 data for separating species.

331 Morphologically, the characters used to differentiate *Eukoenenia* species are mostly
332 restricted to the number of lobules in the lateral organs or the number of setae in different body
333 regions, but the significance of these characters has not been tested phylogenetically—for
334 example, *E. mirabilis* and *E. ferratilis* are very similar morphologically with many somatic traits,
335 considered important for taxonomy, virtually identical (Souza and Ferreira 2011a). However,
336 these two species belong to different clades, reflecting that their differences in genital
337 morphology and chaetotaxy may be better systematic characters than the ones outlined above.
338 Our study thus provides a new framework for adding new sequences and testing the significance
339 of these characters. Additional samples and especially more genera must however be added
340 before we can attempt a taxonomic revision of the higher taxa in Palpigradi.

341

342 **Conclusions**

343 Palpigrades are a poorly understood group of tiny soil arthropods, often found exclusively in
344 caves, and have received little attention from a phylogenetic point of view. Here we were able

345 to amass specimens from different environments (caves and soil) from Australia, Africa, Europe,
346 South America and North America with the aim of generating a molecular phylogenetic
347 hypothesis for the group. The difficulty in obtaining well-preserved material for molecular work
348 is reflected in the large number of specimens that did not yield DNA of enough quality for
349 sequencing, but we were able to propose the first phylogenetic hypothesis of the group based
350 on molecular data to find monophyly of Eukoeneniidae and its division into four main clades,
351 three of these including samples from multiple continents. Given the absence of denser
352 sampling and proper clock calibrations, our data cannot discern whether palpigrades are a very
353 old group that diversified prior to the breakup of Pangaea, or a group of animals that disperses
354 across large geographic distances, as suggested by some widespread species. Long-range
355 dispersal is however difficult to reconcile with the narrow ecological conditions and the facility
356 with which these animals desiccate once removed from their environments.

357

358 **Acknowledgements**

359 Julián Bueno-Villegas and Jesús A. Cruz-López helped with fieldwork in Yucatán; T. Delić and S.
360 Prevorčnik in Slovenia; P. Ľuptáčik and A. Mock in Slovakia; L. Galli and M. Zinni provided
361 samples from Italy; J. Van der Schyff from South Africa; J. Christophoryová from Slovakia; staff
362 from the environmental companies Biota Environmental Sciences and Subterranean Ecology
363 provided samples from Western Australia. Gustavo Hormiga and Nikolaj Scharff kindly provided
364 comments that helped improve this article. This work has been supported by internal funds
365 from the Museum of Comparative Zoology and by NSF grant #1144417 to G. G. and G. Hormiga
366 (Collaborative Research: ARTS: Taxonomy and systematics of selected Neotropical clades of
367 arachnids).

## References


Arabi, J., Judson, M. L., Deharveng, L., Lourenço, W. R., Cruaud, C., and Hassanin, A. (2012). Nucleotide composition of CO1 sequences in Chelicerata (Arthropoda): detecting new mitogenomic rearrangements. *Journal of Molecular Evolution* **74**, 81-95. doi:10.1007/s00239-012-9490-7

Barranco, P., and Harvey, M. S. (2008). The first indigenous palpigrade from Australia: a new species of *Eukoenenia* (Palpigradi : Eukoeneniidae). *Invertebrate Systematics* **22**, 227-233.

Beccaloni, J. (2009). 'Arachnids.' (The Natural History Museum: London.)

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* **17**, 540-552.

Christian, E., Capurro, M., and Galli, L. (2010). Phenology of two syntopic *Eukoenenia* species in a northern Italian forest soil (Arachnida: Palpigradi). *Revue suisse de Zoologie* **117**, 829-834.

Condé, B. (1996). Les palpigrades, 1885-1995: acquisitions et lacunes. *Revue suisse de Zoologie* **hors série**, 87-106.

Delclòs, X., Nei, A., Azar, D., Bechly, G., Dunlop, J. A., Engel, M. S., and Heads, S. W. (2008). The enigmatic Mesozoic insect taxon Chresmodidae (Polyneoptera): New palaeobiological and phylogenetic data, with the description of a new species from the Lower Cretaceous of Brazil. *Neues Jahrbuch für Geologie und Palaontologie-Abhandlungen* **247**, 353-381.

Dimitrov, D., Lopardo, L., Giribet, G., Arnedo, M. A., Álvarez-Padilla, F., and Hormiga, G. (2012). Tangled in a sparse spider web: single origin of orb weavers and their spinning work unravelled by denser taxonomic sampling. *Proceedings of the Royal Society B: Biological Sciences* **279**, 1341-1350. doi:10.1098/rspb.2011.2011

Dunlop, J. A. (2010). Geological history and phylogeny of Chelicerata. *Arthropod Structure & Development* **39**, 124-142. doi:10.1016/j.asd.2010.01.003

395 Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high
396 throughput. *Nucleic Acids Research* **32**, 1792-1797. doi:10.1093/nar/gkh340

397 Edgecombe, G. D., and Giribet, G. (2008). A New Zealand species of the trans-Tasman centipede
398 order Craterostigmomorpha (Arthropoda : Chilopoda) corroborated by molecular
399 evidence. *Invertebrate Systematics* **22**, 1-15. doi:10.1071/Is07036

400 Edgecombe, G. D., and Giribet, G. (2009). Phylogenetics of scutigeromorph centipedes
401 (Myriapoda: Chilopoda) with implications for species delimitation and historical
402 biogeography of the Australian and New Caledonian faunas. *Cladistics* **25**, 406-427.
403 doi:10.1111/j.1096-0031.2009.00253.x

404 Farris, J. S., Albert, V. A., Källersjö, M., Lipscomb, D., and Kluge, A. G. (1996). Parsimony
405 jackknifing outperforms neighbor-joining. *Cladistics* **12**, 99-124.

406 Ferreira, R. L., and Souza, M. F. V. R. (2012). Notes on the behavior of the advanced troglobite
407 *Eukoenenia maquinensis* Souza & Ferreira 2010 (Palpigradi: Eukoeneniidae) and its
408 conservation status. *Speleobiology Notes* **4**, 17-23.

409 Ferreira, R. L., Souza, M. F. V. R., Machado, E. O., and Brescovit, A. D. (2011). Description of a
410 new *Eukoenenia* (Palpigradi: Eukoeneniidae) and *Metagonia* (Araneae: Pholcidae) from
411 Brazilian caves, with notes on their ecological interactions. *The Journal of Arachnology*
412 **39**, 409-419. doi:10.1636/Ha11-03.1

413 Giribet, G. (2005). Generating implied alignments under direct optimization using POY. *Cladistics*
414 **21**, 396-402.

415 Giribet, G. (2007). Efficient tree searches with available algorithms. *Evolutionary Bioinformatics*
416 *Online* **3**, 341-356.

417 Giribet, G., and Edgecombe, G. D. (2013a). The Arthropoda: a phylogenetic framework. In
418 'Arthropod Biology and Evolution – Molecules, Development, Morphology'. (Eds A.
419 Minelli, G. Boxshall and G. Fusco) pp. 17-40. (Springer: Berlin.)

420 Giribet, G., and Edgecombe, G. D. (2013b). Stable phylogenetic patterns in scutigeromorph
421 centipedes (Myriapoda : Chilopoda : Scutigeromorpha): dating the diversification of an

ancient lineage of terrestrial arthropods. *Invertebrate Systematics* **27**, 485-501. doi:10.1071/IS13019

Giribet, G., Edgecombe, G. D., Wheeler, W. C., and Babbitt, C. (2002). Phylogeny and systematic position of Opiliones: a combined analysis of chelicerate relationships using morphological and molecular data. *Cladistics* **18**, 5-70.

Giribet, G., Sharma, P. P., Benavides, L. R., Boyer, S. L., Clouse, R. M., de Bivort, B. L., Dimitrov, D., Kawauchi, G. Y., Murienne, J. Y., and Schwendinger, P. J. (2012). Evolutionary and biogeographical history of an ancient and global group of arachnids (Arachnida: Opiliones: Cyphophthalmi) with a new taxonomic arrangement. *Biological Journal of the Linnean Society* **105**, 92-130. doi:10.1111/J.1095-8312.2011.01774.X

Giribet, G., and Shear, W. A. (2010). The genus *Siro* Latreille, 1796 (Opiliones, Cyphophthalmi, Sironidae), in North America with a phylogenetic analysis based on molecular data and the description of four new species. *Bulletin of the Museum of Comparative Zoology* **160**, 1-33.

Giribet, G., Vogt, L., Pérez González, A., Sharma, P., and Kury, A. B. (2010). A multilocus approach to harvestman (Arachnida: Opiliones) phylogeny with emphasis on biogeography and the systematics of Laniatores. *Cladistics* **26**, 408-437. doi:10.1111/j.1096-0031.2009.00296.x

Goloboff, P. A. (1999). Analyzing large data sets in reasonable times: solutions for composite optima. *Cladistics* **15**, 415-428.

Goloboff, P. A. (2002). Techniques for analyzing large data sets. In 'Techniques in Molecular Systematics and Evolution'. (Eds R. DeSalle, G. Giribet and W. Wheeler) pp. 70-79. (Brikhäuser Verlag: Basel.)

Grassi, B., and Calandruccio, S. (1885). Intorno ad un nuovo aracnide artrogastro (*Koenenia mirabilis*) che crediamo rappresentante d'un nuovo ordine (Microteliphonida). *Naturalista Siciliano* **4**, 127-133, 162-169.

Harvey, M. S. (2002). The neglected cousins: What do we know about the smaller arachnid orders? *The Journal of Arachnology* **30**, 357-372.

450 Harvey, M. S., Stáhlavsky, F., and Theron, P. D. (2006). The distribution of *Eukoenenia mirabilis*
451 (Palpigradi: Eukoeneniidae): a widespread tramp. *Records of the Western Australian*
452 *Museum* **23**, 199-203.

453 Kováč, L., Mock, A., Ľuptáčik, P., and Palacios-Vargas, J. G. (2002). Distribution of *Eukoenenia*
454 *spelaea* (Peyerimhoff, 1902) (Arachnida, Palpigradida) in the Western Carpathians with
455 remarks on its biology and behaviour. In 'Studies on Soil Fauna in Central Europe'. (Eds K.
456 Tajovský, V. Balík and V. Pižl) pp. 93-99: České Budějovice.)

457 Král, J., Kováč, L., Št'ahlavský, F., Lonský, P., and L'uptácik, P. (2008). The first karyotype study in
458 palpigrades, a primitive order of arachnids (Arachnida: Palpigradi). *Genetica* **134**, 79-87.

459 Linton, E. W. (2005). MacGDE: Genetic Data Environment for MacOS X. (Software available at
460 http://www.msu.edu/~lintone/macgde/)

461 Mallatt, J., and Giribet, G. (2006). Further use of nearly complete 28S and 18S rRNA genes to
462 classify Ecdysozoa: 37 more arthropods and a kinorhynch. *Molecular Phylogenetics and*
463 *Evolution* **40**, 772-794. doi:10.1016/J.Ympev.2006.04.021

464 Mayoral, J. G., and Barranco, P. (2013). Rediscovery of the troglobious palpigrade *Eukoenenia*
465 *draco* (Peyerimhoff 1906) (Palpigradi: Eukoeneniidae), with notes on the adaptations to
466 a cave-dwelling life. *Zootaxa* **3635**, 174-184. doi:10.11646/zootaxa.3635.2.5

467 Miller, M. A., Pfeiffer, W., and Schwartz, T. (2010). Creating the CIPRES Science Gateway for
468 Inference of Large Phylogenetic Trees. In 'Proceedings of the Gateway Computing
469 Environments Workshop (GCE)'. New Orleans pp. 1-8

470 Montaño Moreno, H. (2008). Revisión taxonómica de los palpígrados (Arachnida: Palpigradi) de
471 México. Masters thesis, Universidad Nacional Autónoma de México.

472 Montaño-Moreno, H. (2012). Redescripción de *Eukoenenia hanseni* (Arachnida: Palpigradi) y
473 descripción de una nueva especie de palpígrado de México. *Revista Ibérica de*
474 *Aracnología* **20**, 1-15.

475 Murienne, J., Harvey, M. S., and Giribet, G. (2008). First molecular phylogeny of the major clades
476 of Pseudoscorpiones (Arthropoda: Chelicerata). *Molecular Phylogenetics and Evolution*
477 **49**, 170-184. doi:10.1016/j.ympev.2008.06.002

478 Nixon, K. C. (1999). The Parsimony Ratchet, a new method for rapid parsimony analysis.
479 *Cladistics* **15**, 407-414.

480 Pepato, A. R., da Rocha, C. E., and Dunlop, J. A. (2010). Phylogenetic position of the acariform
481 mites: sensitivity to homology assessment under total evidence. *BMC Evolutionary*
482 *Biology* **10**, 235. doi:10.1186/1471-2148-10-235

483 Prendini, L. (2011). Order Palpigradi Thorell, 1888 (In: Animal biodiversity: An outline of higher-
484 level classification and survey of taxonomic richness). *Zootaxa* **3148**, 121.

485 Regier, J. C., Shultz, J. W., Zwick, A., Hussey, A., Ball, B., Wetzer, R., Martin, J. W., and
486 Cunningham, C. W. (2010). Arthropod relationships revealed by phylogenomic analysis
487 of nuclear protein-coding sequences. *Nature* **463**, 1079-1083. doi:10.1038/nature08742

488 Rowland, J. M., and Sissom, W. (1980). Report on a fossil palpigrade from the Tertiary of Arizona,
489 and a review of the morphology and systematics of the order (Arachnida, Palpigradida).
490 *The Journal of Arachnology* **8**, 69-86.

491 Sharma, P. P., Vahtera, V., Kawauchi, G. Y., and Giribet, G. (2011). Running $_{w}$ILD: The case for
492 exploring mixed parameter sets in sensitivity analysis. *Cladistics* **27**, 538-549.
493 doi:10.1111/j.1096-0031.2010.00345.x

494 Shultz, J. W. (2007). A phylogenetic analysis of the arachnid orders based on morphological
495 characters. *Zoological Journal of the Linnean Society* **150**, 221-265.

496 Smrž, J., Kováč, Ľ., Mikeš, J., and Lukešová, A. (2013). Microwhip scorpions (Palpigradi) feed on
497 heterotrophic Cyanobacteria in Slovak caves – A curiosity among Arachnida. *PLoS ONE* **8**,
498 e75989. doi:10.1371/journal.pone.0075989

499 Souza, M. F. V. R., and Ferreira, R. L. (2010). *Eukoenenia* (Palpigradi: Eukoeneniidae) in Brazilian
500 caves with the first troglobiotic palpigrade from South America. *The Journal of*
501 *Arachnology* **38**, 415-424.

502 Souza, M. F. V. R., and Ferreira, R. L. (2011a). A new species of *Eukoenenia* (Palpigradi:
503 Eukoeneniidae) from Brazilian iron caves. *Zootaxa* **2886**, 31-38.

504 Souza, M. F. V. R., and Ferreira, R. L. (2011b). A new troglobitic *Eukoenenia* (Palpigradi:
505 Eukoeneniidae) from Brazil. *Journal of Arachnology* **39**, 185-188. doi:10.1636/Ha10-43.1

506 Souza, M. F. V. R., and Ferreira, R. L. (2012a). *Eukoenenia virgemdalapa* (Palpigradi:
507 Eukoeneniidae): a new troglobitic palpigrade from Brazil. *Zootaxa* **3295**, 59-64.

508 Souza, M. F. V. R., and Ferreira, R. L. (2012b). A new highly troglomorphic species of *Eukoenenia*
509 (Palpigradi: Eukoeneniidae) from tropical Brazil. *The Journal of Arachnology* **40**, 151-158.

510 Souza, M. F. V. R., and Ferreira, R. L. (2013). Two new species of the enigmatic Leptokoenenia
511 (Eukoeneniidae: Palpigradi) from Brazil: First record of the genus outside intertidal
512 environments. *PLoS ONE* **8**, e77840. doi:10.1371/journal.pone.0077840

513 Stamatakis, A., Hoover, P., and Rougemont, J. (2008a). A rapid bootstrap algorithm for the
514 RAxML Web servers. *Systematic Biology* **57**, 758-771. doi:10.1080/10635150802429642

515 Stamatakis, A. P., Meier, H., and Ludwig, T. (2008b). RAxML: A parallel program for phylogenetic
516 tree inference.

517 Talavera, G., and Castresana, J. (2007). Improvement of phylogenies after removing divergent
518 and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology*
519 **56**, 564-577. doi:10.1080/10635150701472164

520 Vaidya, G., Lohman, D. J., and Meier, R. (2011). SequenceMatrix: concatenation software for the
521 fast assembly of multi-gene datasets with character set and codon information.
522 *Cladistics* **27**, 171-180. doi:10.1111/j.1096-0031.2010.00329.x

523 Varón, A., Lucaroni, N., Hong, L., and Wheeler, W. C. (2012). POY 5.0.0. (American Museum of
524 Natural History. http://research.amnh.org/scicomp: New York)

525 Vélez, S., Mesibov, R., and Giribet, G. (2012). Biogeography in a continental island: population
526 structure of the relict endemic centipede *Craterostigmus tasmanianus* (Chilopoda,
527 Craterostigmomorpha) in Tasmania using 16S rRNA and COI. *Journal of Heredity* **103**, 80-
528 91. doi:10.1093/jhered/esr110

529 Wheeler, W. C. (1995). Sequence alignment, parameter sensitivity, and the phylogenetic analysis
530 of molecular data. *Systematic Biology* **44**, 321-331.

Wheeler, W. (1996). Optimization alignment: the end of multiple sequence alignment in phylogenetics? *Cladistics* **12**, 1-9.

Wheeler, W. C. (2003). Implied alignment: a synapomorphy-based multiple-sequence alignment method and its use in cladogram search. *Cladistics* **19**, 261-268.

Wheeler, W. C., and Hayashi, C. Y. (1998). The phylogeny of the extant chelicerate orders. *Cladistics* **14**, 173-192.

Whiting, M. F., Carpenter, J. M., Wheeler, Q. D., and Wheeler, W. C. (1997). The Strepsiptera problem: phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. *Systematic Biology* **46**, 1-68.

Zagmajster, M., and Kováč, Ľ. (2006). Distribution of palpigrades (Arachnida, Palpigradi) in Slovenia with a new record of *Eukoenenia austriaca* (Hansen, 1926). *Natura Sloveniae* **8**, 23-31.

544 **Fig. 1.** Photographs of (A) *Eukoenenia spelaea*, Ardovská Cave (Slovak Karst, Slovakia),
545 photographed by Ľ. Kováč & V. Kóňa; (B) *Prokoenenia wheeleri*, Austin (Texas, USA),
546 photographed by L. McCutchen; (C) *Eukoenenia mirabilis*, flagellum, segments 1 to 10; (D)
547 *Eukoenenia bonadonai*, male genital lobes; (E) *E. bonadonai*, female genital lobes; (F) *E.*
548 *bonadonai*, mouth cone and chelicerae (C-D photographed by E. Christian).

549

550 **Fig. 2.** Left: Optimal tree at 10,408 weighted steps obtained from the direct optimization
551 analysis under parameter set 3211 of the combined analysis of the three genes. Numbers on
552 branches indicate jackknife support values. Navajo rugs are shown in selected nodes; Black
553 square indicates monophyly, white square non-monophyly. Specific parameter sets or analyses
554 indicated in the figure. Numerals indicate parameter set under parsimony direct optimization; IA
555 (ML analysis using implied alignment under parameter set 3211); IAg (Idem, Gblocked); MSA
556 (ML analysis of the MUSCLE multiple sequence alignment); MSAg (Idem, Gblocked). Clades I to
557 IV are indicated.

558

559 **Fig. 3.** Optimal maximum likelihood tree (-Ln*L* = -24955.690470) of the combined data set using
560 the MUSCLE multiple sequence alignment trimmed with Gblocks. Numbers on nodes indicate
561 bootstrap support values.

**Table 1. Palpigrade specimens, accession numbers, collecting information and amplified loci with GenBank accession numbers**

IZ: Department of Invertebrate Zoology, Museum of Comparative Zoology, Cambridge; DNA: MCZ DNA collection; WAM: Western Australian Museum, Perth; MNHN: Muséum national d'histoire Naturelle, Paris. A dash (-) indicates a missing amplicon. New sequences are KF823823 to KF823883

| | | | | | 18S rRNA | | | 28S rRNA | | COI |
|---|---|---|---|---|---|---|---|---|---|---|
| | ***MCZ No.*** | | ***Country*** | ***a*** | ***b*** | ***c*** | ***a*** | ***b*** | ***c*** | |
| *Prokoenenia wheeleri* | IZ-134477 | DNA107078 | Texas, USA | KF823823 | KF823823 | KF823823 | KF823848 | KF823848 | KF823848 | KF823874 |
| *Eukoenenia austriaca* | IZ-19349 | - | Slovenia | KF823824 | KF823824 | KF823824 | KF823849 | KF823849 | KF823849 | - |
| *Eukoenenia bonadonai* | IZ-19340 | - | Italy | KF823825 | KF823825 | KF823825 | KF823850 | KF823850 | KF823850 | - |
| *Eukoenenia ferratilis* | IZ-127609 | - | Brazil | KF823826 | KF823826 | KF823826 | KF823851 | KF823851 | KF823851 | - |
| *Eukoenenia* cf. *ferratilis* | - | GenBank | | HM070336 | HM070336 | HM070336 | HM070299 | HM070299 | HM070299 | - |
| *Eukoenenia florenciae* | IZ-19351 | - | Slovakia | KF823827 | KF823827 | KF823827 | KF823852 | KF823852 | KF823852 | KF823875 |
| *Eukoenenia* cf. *florenciae* | IZ-19343 | - | Brazil | KF823828 | KF823828 | KF823828 | KF823853 | KF823853 | KF823853 | - |
| *Eukoenenia mirabilis* | IZ-127901 | - | Italy | KF823829 | KF823829 | KF823829 | KF823854 | KF823854 | KF823854 | KF823876 |
| *Eukoenenia mirabilis* | IZ-127902 | - | Italy | KF823830 | KF823830 | KF823830 | KF823855 | KF823855 | KF823855 | KF823877 |
| *Eukoenenia mirabilis* | IZ-16117 | - | Australia | KF823831 | KF823831 | KF823831 | KF823856 | KF823856 | KF823856 | - |
| *Eukoenenia spelaea* | IZ-135126 | DNA106786 | Slovakia | - | KF823832 | KF823832 | KF823857 | KF823857 | KF823857 | - |
| *Eukoenenia spelaea* | IZ-19346 | - | Slovenia | KF823833 | KF823833 | KF823833 | KF823858 | KF823858 | KF823858 | KF823878 |
| *Eukoenenia spelaea* | IZ-19347 | - | Slovenia | KF823834 | KF823834 | KF823834 | KF823859 | KF823859 | KF823859 | - |
| *Eukoenenia spelaea hauseri* | IZ-19348 | - | Slovenia | KF823835 | KF823835 | KF823835 | KF823860 | KF823860 | KF823860 | - |
| *Eukoenenia strinatii* | IZ-19341 | - | Italy | KF823836 | KF823836 | KF823836 | KF823861 | KF823861 | KF823861 | - |
| *Eukoenenia* sp. | IZ-19350 | - | Slovenia | KF823837 | KF823837 | KF823837 | KF823862 | KF823862 | KF823862 | KF823879 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *Eukoenenia* sp. | - | DNA100456.1 | South Africa | AF207648 | AF207648 | AF207648 | - | AF207653 | - | - |
| *Eukoenenia* sp. | - | DNA100456.2 | South Africa | KF823838 | - | KF823839 | - | KF823863 | - | - |
| *Eukoenenia* sp. | IZ-134549 | DNA107079 | USA | KF823840 | KF823840 | KF823840 | KF823864 | KF823864 | KF823864 | KF823880 |
| *Eukoenenia* sp. | IZ-127598.1 | - | Mexico | KF823841 | KF823841 | KF823841 | KF823865 | KF823865 | KF823865 | KF823881 |
| *Eukoenenia* sp. | IZ-127598.2 | - | Mexico | KF823842 | KF823842 | KF823842 | KF823866 | KF823866 | KF823866 | KF823882 |
| *Eukoenenia* sp. | IZ-128499 | - | Mexico | KF823843 | KF823843 | KF823843 | KF823867 | KF823867 | KF823867 | KF823883 |
| *Eukoenenia* sp. | IZ-136274 | - | Mexico | KF823844 | - | KF823844 | KF823868 | KF823868 | KF823868 | - |
| *Eukoenenia* sp. | IZ-127636 | WAM T81111 | Australia | - | - | - | - | KF823869 | - | - |
| *Eukoenenia* sp. | IZ-127639 | WAM T116012 | Australia | KF823845 | KF823845 | KF823845 | - | KF823870 | KF823870 | - |
| *Eukoenenia* sp. | IZ-127640 | WAM T111422 | Australia | - | - | - | - | KF823871 | - | - |
| *Eukoenenia* sp. | IZ-127643 | - | Australia | KF823846 | KF823846 | KF823846 | - | KF823872 | KF823872 | - |
| *Eukoenenia* sp.n. | IZ-19345 | - | Brazil | KF823847 | KF823847 | KF823847 | - | KF823873 | KF823873 | - |
| Palpigradi sp. | - | MNHN-JAA76 | | JN018286.1 | JN018286.1 | JN018286.1 | JN018383.1 | JN018383.1 | JN018383.1 | JN018169.1 |

**Table 2. Outgroup sampling with GenBank accession numbers**

| | | 18S rRNA | 28S rRNA | COI |
|---|---|---|---|---|
| *Anoplodactylus portus* | Pycnogonida | AY859551 | AY859550 | GQ912859 |
| *Limulus polyphemus* | Xiphosura | U91490 | AF212167 | AF216203 |
| *Pandinus imperator* | Scorpiones | AY210831 | AY210830 | AY156582 |
| *Metasiro americanus* | Opiliones | DQ825542 | DQ825595 | DQ825645 |
| *Calocheiridius termitophilus* | Pseudoscorpiones | AY859559 | AY859558 | EU559544 |
| *Dermacentor* sp. | Acari | Z74480 | AY859582 | - |
| *Eremobates* sp. | Solifugae | AY859573 | AY859572 | - |
| *Mastigoproctus giganteus* | Uropygi | AF005446 | AY859587 | JN018215 |

**Table 3. Result of the POY timed searches (search) and improvement after each round of SATF for the six explored parameter sets**

| | 1 | SATF2 | SATF3 |
|---|---|---|---|
| **111** | 6520 | 6520 | 6520 |
| **121** | 10076 | 10076 | 10076 |
| **211** | 7543 | 7543 | 7543 |
| **221** | 11851 | 11851 | 11851 |
| **3211** | 10408 | 10408 | 10408 |
| **3221** | 13526 | 13526 | 13526 |

**Table 4. Number of weighted steps for each data partition, the combination of them (MOL) and $_{w}$ILD value**

The optimal parameter set is indicated in italics

| | 18S | 28S | COI | MOL | wILD |
|---|---|---|---|---|---|
| **111** | 1125 | 3967 | 1354 | 6520 | 0.01135 |
| **121** | 1655 | 6272 | 2051 | 10076 | 0.00973 |
| **211** | 1246 | 4840 | 1381 | 7543 | 0.01008 |
| **221** | 1867 | 7780 | 2080 | 11851 | 0.01046 |
| ***3211*** | *1704* | *6535* | *2074* | *10408* | *0.00913* |
| **3221** | 2314 | 8305 | 2777 | 13526 | 0.00961 |

**Table 5. Length of each data partition (28S rRNA is divided into three amplicons) and total length of alignment**

IA (121) is for implied alignment under parameter set 121; IA+Gb is for implied alignment trimmed with Gblocks; Muscle is for MUSCLE multiple sequence alignment; Muscle+Gb is for multiple sequence alignment trimmed with Gblocks

| | 18S | 28Sa | 28Sbc | COI | TOTAL |
|---|---|---|---|---|---|
| **Unaligned** | 1760-1805 | 832-873 | 1265-1347 | 654-657 | |
| **IA (3211)** | 1860 | 1323 | 1555 | 669 | 5407 |
| **IA+Gb** | 1676 | 378 | 1162 | 626 | 3842 |
| **Muscle** | 1818 | 1046 | 1409 | 663 | 4936 |
| **Muscle+Gb** | 1695 | 609 | 1212 | 636 | 4152 |